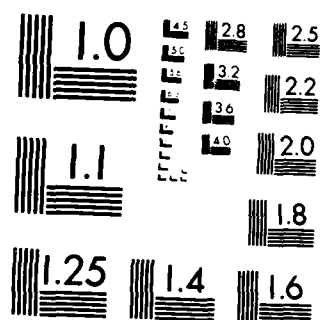


EVIDENCE ACCUMULATION FOR SPATIAL REASONING(U) MARYLAND 1/  
UNIV COLLEGE PARK CENTER FOR AUTOMATION RESEARCH  
T MATSUYAMA ET AL. MAR 84 CAR-TR-54 AFOSR-TR-84-0468  
540608-02-C 8882

F/G 14/5

Ni

END  
DATE  
FILMED  
7 84  
DTIC



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS 1963-A

AD-A142 174

CAR-TR-54  
CS-TR-1381

F-49620-83-C-0082  
March 1984

**Evidence Accumulation for Spatial Reasoning**

**Takashi MATSUYAMA\***  
**Vincent Shang-Shouq HWANG**  
**Larry S. DAVIS**

**Center for Automation Research**  
**University of Maryland**  
**College Park, MD 20742**

CAR-TR-54  
CS-TR-1381

F-49620-83-C-0082  
March 1984

**Evidence Accumulation for Spatial Reasoning**

**Takashi MATSUYAMA\***  
**Vincent Shang-Shouq HWANG**  
**Larry S. DAVIS**

**Center for Automation Research**  
**University of Maryland**  
**College Park, MD 20742**

**ABSTRACT**

*This paper describes the evidence accumulation process of an image understanding system first described in [1], which enables the system to perform top-down(goal-oriented) picture processing as well as bottom-up verification of consistent spatial relations among objects.*

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFOSR)  
NOTICE OF REPLY TO THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH  
THIS DOCUMENT IS UNCLASSIFIED  
DATE 11-12-83 BY 1045  
MATTHEW J. KLEIN  
Chief, Technical Information Division

**\*Permanent address:**

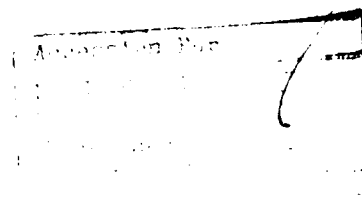
**Department of Electrical Engineering**  
**Kyoto University, Sakyo, Kyoto, Japan**

The support of the U.S. Air Force Office of Scientific Research under Contract F-49620-83-C-0082 is gratefully acknowledged.

## 1. Introduction

In a previous report[1], we described the organization of an aerial image analysis system. There are three levels of representation and control in that system: A High Level Expert(HLE) that utilizes a symbolic hierarchical model for the possible spatial organization of objects in the image to build partial, local interpretations of the image and to determine where to further analyze the image and what analyses to perform; a Model Selection Expert(MSE) that determines, on the basis of contextual information provided by the HLE, the most promising appearance descriptions to use in searching for objects and structures in the image; and a Low Level Vision Expert(LLVE) that finds pictorial entities that satisfy these appearance descriptions by selecting image processing methods to find the appropriate entities.

Our emphasis has been on the High Level Expert, which is based on a general method of "evidence accumulation" to perform flexible spatial reasoning. This paper contains a detailed description of our evidence accumulation process and its associated consistency checking process.



## 2. Motivation

In general, two different types of information can be used to interpret a pictorial entity: its intrinsic properties(size, shape, color etc.) and its relations to other entities. Our primary interest is the representation of geometric relations among objects and their utilization for image interpretation. This is especially important in recognition of man-made objects. Moreover, although shape can often be regarded as an intrinsic object property, a complex shape is often described structurally in terms of geometric relations among its components. Thus shape recognition often requires spatial analysis.

Let  $REL(O1, O2)$  denote a binary geometric relation between two classes of objects,  $O1$  and  $O2$ . This relation can be used as a constraint to recognize objects from these two classes by first extracting pictorial entities which satisfy the intrinsic properties of  $O1$  and  $O2$ , and then checking that the geometric relation is satisfied by these candidate objects(Figure 1). In this *bottom-up* recognition scheme, analysis based on geometric relations cannot be performed until pictorial entities corresponding to objects are extracted.

In general, however, some of the correct pictorial entities often fail to be extracted by the initial image segmentation. So one must, additionally, incorporate *top-down control* to find pictorial entities missed by the initial segmentation. Such top-down processes use geometric relations to predict the locations of missing objects, as in the system described by Selfridge[2].

It is, of course, generally accepted that image understanding systems should incorporate both bottom-up and top-down analyses. As noted above, the use of geometric relations is very different in the two analysis processes: consistency verification in bottom-up analysis and hypothesis generation in top-down analysis. An important characteristic of our evidence accumulation method is that it enables the system to integrate both bottom-up and top-down processes into a single flexible spatial reasoning process. As will be described later, the system first establishes local environments. Then, either bottom-up or top-down processes are activated depending on the nature of the local environment. The following sections describe the concepts and characteristics of this process.

### 3. Representation of Geometric Relations and Hypothesis Formation

#### 3.1. Functional Representation of Relations

A relation  $REL(O1, O2)$  ( $O1$  and  $O2$  are object classes) is represented using two functional expressions:

$$O1 = f(O2) \text{ and } O2 = g(O1).$$

Given an instance of  $O2$ , say  $r$ , function  $f$  maps it into a description of an instance of  $O1$ ,  $f(r)$ , which satisfies the geometric relation,  $REL$ , with  $r$ . The analogous interpretation holds for the other function  $g$ .

In our system, knowledge about a class of objects is represented by a frame[3], and a slot in that frame is used to store a function such as  $f$  or  $g$ . The function is represented by a computational procedure(which produces the description of the related instance) and a set of conditions to specify when that function can be activated. Whenever an *instance* of an object is created, and the conditions are satisfied, the function is applied to the instance to generate a *hypothesis* (expectation) for another object which would, if found, satisfy the geometric relation with the original instance. The function can use any properties of the instance to create the hypothesis.

A hypothesis is associated with a prediction area where the related object instance may be located(Figure 2). In addition to this area specification, a set of constraints on the target instance is associated with the hypothesis. Figure



3 shows the description of a road hypothesis. All hypotheses and instances are stored in a common database(the *iconic* database) where *accumulation of evidence* (i.e., recognition of overlapping sets of consistent hypotheses and instances) is performed. Similar ideas have been proposed to solve spatial layout problems[4] and to answer queries about map information[5].

### **3.2. Spatial Relations, Part-Whole Relations, and A-Kind-Of Relations**

Two types of geometric relations are used in our system: "spatial relation"(SP) and "part-whole relation"(PW). These two types of relations are used differently by the system. The PW relations specify AND/OR hierarchies which represent objects with complex internal structure. The SP relations represent geometric and topological relations between objects. In addition, "A-kind-of relations"(AKO) are used to construct object specialization hierarchies.

There are several restrictions on the usage of these types of relations. A hierarchy defined by the PW relation must be a tree structure. Although SP relations can be established across objects in different PW hierarchies, an object cannot have an SP relation with another object in the same PW hierarchy, nor can it establish multiple SP relations to any other PW hierarchy. These restrictions were adopted to avoid redundant generation of hypotheses.

Consider the knowledge representations shown in Figures 4(a) and (b). If object A had an SP relation to object B in the same part-whole hierarchy(Figure 4(a)), there would be two paths from object A to generate a hypothesis of object B: one by the SP relation and the other by the PW relation. This means that if an instance of object A were constructed, two hypotheses for object B would be generated from the same instance. The same argument holds in the case shown in Figure 4(b). Figure 4(c) shows a circular path consisting of SP relations between objects A, B, and C. This is allowed since no redundant hypotheses are formed.

Hypothesis generation by an SP relation is done as explained above, i.e., when an object is instantiated and the set of conditions needed to generate a hypothesis are satisfied, then the function associated with the SP relation is activated to produce an expectation area and an associated set of constraints for a target object. Although, syntactically, SP relations represent binary relations, it is possible to use them to represent n-ary relations. For example, a left eye can create a hypothesis for a nose, and can use the known location of a potential right eye to generate the nose hypothesis.

The system uses PW relations both to group parts into a whole and to predict missing parts. If an instantiated object corresponds to a leaf node in the PW hierarchy, then it can directly *instantiate* (again, if prespecified conditions hold) its parent node through the PW relation(Figure 5).

Objects at the leaves of PW hierarchies are instantiated first, since they correspond directly to low-level image structures. The presence of a higher level object is represented by an instantiated PW hierarchy. The parent may then *hypothesize* the presence of other missing object parts. For computational simplicity, there are no hypotheses generated between siblings in the PW hierarchy.

## 4. Combining Evidence

### 4.1. The Interpretation Cycle of the High Level Expert

Figure 6 shows the organization of the entire system. The High Level Expert iterates the following steps.

- (1) Each instance of an object generates hypotheses about related objects using functions stored in the object model(frame).
- (2) All pieces of evidence(both instances and hypotheses) are stored in a common database(iconic database). They are represented using an iconic data structure which associates highly structured symbolic descriptions of the instances and hypotheses with regions in a two-dimensional array.
- (3) Pieces of evidence are combined to establish *situations*. A situation consists of *consistent* pieces of evidence.
- (4) Focus of attention : since there are many situations, the most reliable situation is selected.
- (5) The selected situation is *resolved*, which results either in verification of predictions on the basis of previously detected/constructed image structures or in top-down image processing to detect missing objects.

The system also has two additional processes:

- (1) Instantiation of objects at the very beginning of interpretation

This process is performed by the Model Selection Expert which searches for object models that have simple appearances, and directs the Low Level Vision

Expert to detect pictorial entities which satisfy the appearances. The instances constructed by this process are seeds for reasoning by the High Level Expert.

#### (2) Selection of the maximum consistent interpretation

During the analysis by the High Level Expert, inconsistent pieces of evidence may be constructed. The High Level Expert maintains all possible interpretations throughout the search process until no further changes are made in the iconic database. A final interpretation then selects the maximal consistent interpretation.

The following subsections provide detailed discussion of the operation of the High Level Expert.

### 4.2. Overview

Given a set of instances of objects, each of them activates functions to generate hypotheses about related objects. Each instance and hypothesis is represented as a region in the iconic data structure. Suppose instance  $s$  creates hypothesis  $f(s)$  (based on relation  $R$ ) for object class  $O1$ , which overlaps with an instance of  $O1$ ,  $t$  (Figure 7(a)). If the set of constraints associated with  $f(s)$  is satisfied by  $t$ , these two pieces of evidence are combined to form what we call a *situation*. The more pieces of evidence that are combined, the more reliable the situation becomes. The High Level Expert unifies  $f(s)$  and  $t$ , and establishes the relation  $R$  from  $s$  to  $t$  as the result of *resolving* the situation.

On the other hand, a situation may consist of overlapping *hypotheses*, if their constraints are consistent(Figure 7(b)). Then their unification leads the expert to search for an instance of the required object in the image. The High Level Expert asks the Model Selection Expert to detect the instance, which in turn activates the Low Level Vision Expert. If the instance is detected, it is inserted into the database. Hypothesis generation by the newly detected instance is performed at the next interpretation cycle.

### 4.3. Handling PW relations

Additional complications arise from resolving situations involving instances generated via PW relations. Suppose  $s$  is an instance of an object corresponding to a leaf node in a PW hierarchy(Figure 8(a)). As described above, it may instantiate its parent object. Let  $p$  denote this instance. Then  $p$  generates a hypothesis for a missing part,  $f(p)$ . If there is already an instance corresponding to the missing part, say  $t$ ,  $f(p)$  and  $t$  will be unified, and a part-whole relation will be established between  $p$  and  $t$ . However, since  $t$  is also an instance, it may also have instantiated its parent object. Let  $u$  denote this instance. As the result of the unification, instance  $t$  has two parent instances,  $p$  and  $u$ . This leads the High Level Reasoning Expert to another unification. The expert examines  $p$  and  $u$ , and if they are consistent, it unifies them(Figure 8(b)). This unification may trigger still another unification for higher level instances in the hierarchy. Note that after the unification, instance  $p$  can use properties of  $r$  and  $t$  to generate hypotheses for other part

objects whose geometric properties could not previously be specified due to a lack of sufficient information.

If the two parent instances( $p$  and  $u$ ) were found not to be consistent, the expert records such mutually conflicting interpretations, and will perform reasoning independently based on each interpretation. The process of reasoning with alternative interpretations is not described in detail in this paper.

There can be a still more complicated situation created by a PW relation. As shown in Figure 9(a), suppose the grandparent object has also been instantiated by an instance of a leaf object,  $r$ . Let  $p$  and  $q$  denote instances of the parent and grandparent objects, respectively;  $q$  as well as  $p$  generates hypotheses for its missing parts, say  $f(q)$ . Suppose that  $f(q)$  itself has parts and one of them has already been instantiated. Let  $s$  denote that instance. Then, if instances  $r$  and  $s$  are really parts of the same object, regions of  $f(q)$  and  $s$  will overlap with each other and will be consistent.(A detailed discussion of consistency will be given in the next subsection.) In this case, the system first constructs a situation based on the intersection of  $f(q)$  and  $s$ , even if their description levels in the PW hierarchy are different, and then unifies  $f(q)$  and  $t$ (the parent instance of  $s$ ). Note that instance  $t$  cannot intersect with  $f(q)$  directly since no iconic region is associated with  $t$  in the database. As a result,  $r$ ,  $p$ ,  $q$ ,  $s$ , and  $t$  are organized into one hierarchical structure(Figure 9(b)). If, as shown in Figure 9(c), the levels of  $f(q)$  and  $t$  in the hierarchy are different(in Figure 9(b), they are at the same level), a series of parent objects

are instantiated from instance  $s$ .

#### **4.4. Forming Consistent Situations**

Consistent pieces of evidence from different sources are combined into situations. The consistency among pieces of evidence is based on:

- (1) prediction areas of hypotheses
- (2) object categories of evidence
- (3) constraints imposed on properties of hypotheses and instances
- (4) relations among sources of evidence

These criteria are discussed in the next four subsections.

##### **4.4.1. Intersections of Prediction Areas**

Figure 10(a) shows all intersections formed from pieces of evidence  $E_1$ ,  $E_2$ ,  $E_3$ , and  $E_4$ . A partial ordering on intersections can be constructed on the basis of region containment. Intersection  $OP_1$  is less than  $OP_2$  if region  $OP_1$  is contained in region  $OP_2$ . Figure 10(b) shows the lattice representing the intersection in Figure 10(a). Each intersection consists of some set of hypotheses and instance. Situations are only formed among intersecting pieces of evidence.

##### **4.4.2. Object Categories of Evidence**

In our domain, some pairs of objects cannot occupy the same location in an image. For instance, a region cannot be interpreted as both house and road



at the same time(although it could be interpreted both as road and shadow). Pairs of frames representing object classes which cannot occupy the same region are linked with an *in-conflict-with* relation.

Let OP be the intersection arising from evidence  $\{E1, E2\}$  and let OBJ1 and OBJ2 denote the object categories of E1 and E2, respectively. If OBJ1 and OBJ2 are linked by an *in-conflict-with* relation, then E1 and E2 are said to be conflicting, and OP is removed from the lattice. The removal of OP is propagated through the lattice, and any intersections contained in OP are also removed, since they must also have arisen from conflicting evidence. To find all conflicting intersections, it is clearly sufficient to examine all intersections containing only a pair of pieces of evidence and then to propagate the results through the lattice.

In the above case, if both E1 and E2 are *instances*, the High Level Reasoning Expert records them as conflicting and use that fact to establish the inconsistency of situations containing hypotheses generated by conflicting instances. (See Section 4.4.4.)

A shortcoming of our approach to evidence accumulation is that negative sources of evidence are not considered in assessing the strength of a situation. For example, in medical diagnosis, some measurements are used to deny the possibility of certain classes of diseases. Incorporation of sources of negative evidence is an important issue for future research.

#### 4.4.3. Constraint Consistency

After eliminating all conflicting intersections, the remaining intersections are checked to determine if their associated sets of constraints are consistent. Let E1 and E2 denote the non-conflicting evidence under consideration. One of the following conditions must hold:

- (1) The object categories of E1 and E2 are the same,
- (2) there is a path between the two categories consisting of PW relations,
- or
- (3) one piece of evidence is a subcategory of the other, according to the specialization/generalization hierarchy.

In the second case, since the names of the attributes used in the constraints associated with E1 and E2 may be different, they cannot, in general, be directly compared. Suppose the object category of E1 is at a higher level in the hierarchy than that of E2. The constraints associated with E2 are translated into those for the object category of E1 by using part-whole/a-kind-of relations. Then the translated constraints are compared with those associated with E1.

Figure 11 illustrates the translation of constraints using PW relations. Constraint C1 on a road piece object is translated into constraint C2 on a road object. Currently, this translation is done simply by rewriting the

attributes(slot names) of C1 into appropriate attributes(slot names) of C2 using a "slot name translation table" for the PW relation(Figure 11.b).

The properties and/or constraints associated with both pieces of evidence must be consistent. Both constraints associated with a hypothesis and properties associated with an instance are represented by sets of linear inequalities in one variable. A simple constraint manipulation system is used to check the consistency between the sets of inequalities by generating the solution space(also represented by inequalities) to the intersection of sets. If this solution space is empty, then the constraints are inconsistent. If C1 are the constraints for E1, C2 for E2, and C for O, the object category to which both E1 and E2 belong, then we must check that

$$(C1 \cap C2) \text{ and } C \neq \emptyset$$

We do this by first computing  $C3 = (C1 \cap C2)$ , and if this is non-empty, finally computing C3 and C.

#### 4.4.4. Relations Between Sources of Evidence

The *sources* of accumulated evidence about a situation must not be conflicting. Let S1 and S2 denote the source evidence of E1 and E2, respectively. If a piece of evidence is a hypothesis, its source evidence is the instance which generated the hypothesis. An instance is the source evidence for itself. It is possible that S1 and S2 are mutually conflicting(Figure 12), but that E1 and E2 themselves are consistent. In such a case, we do not combine E1 and

E2 into a situation; analysis based on such conflicting interpretations is performed independently.

#### **4.5. Focus of Attention**

After examining the consistency among evidence, we next evaluate the reliability of each consistent situation by summing numerical reliability measures for each piece of evidence, and select the most reliable one for further analysis. This is the focus of attention mechanism.

##### **4.5.1. Controlling the Intermediate Interpretation Process**

Recall that there are two different types of evidence in our system: instances and hypotheses. It is possible to control the direction of the interpretation process by assigning different reliabilities to them.

If a higher reliability is assigned to an instance than to a hypothesis, a situation including an instance tends to be selected as the most reliable one rather than one consisting only of hypotheses. Therefore the system first builds partial interpretations by establishing relations among instances before trying to perform top-down picture processing.

## 5. Resolving a Situation

As described in Section 4.2, one of two actions is taken in order to resolve a situation: confirm relations between instances or activate top-down analysis.

How a situation is resolved depends on the nature of its constituent evidence. If the pieces of evidence are all hypotheses, then a *composite* hypothesis is constructed for transmittal to the MSE, and any instance extracted from the image is then examined by the source instances of those hypotheses. If a situation includes both hypotheses and instances, then the instances are, in turn, examined by the sources of the hypotheses, and if none satisfy the hypotheses, then a composite hypothesis can, in turn, be transmitted to the MSE.

### 5.1. Resolution Process

The system provides a description of its proposed resolution to a situation to all instances involved in that situation. Each instance then evaluates the proposed solution according to its specific expectations.

In what follows, the process of resolving a situation is illustrated by the example shown in Figure 13. Suppose the consistency reasoner selected the overlapping region between two hypotheses generated from two road-piece instances RP1 and RP2(Figure 13(a)). In the symbolic data structure, RP1 and RP2 are linked to their parent road instances RD1 and RD2 by PW rela-

tions, respectively. The hypotheses for adjacent road pieces have been generated by these parent instances.

Since this situation consists only of hypotheses, the system activates top-down analysis to find a road piece in the over-lapping region. This request is issued to the Model Selection Expert together with the supporting evidence(i.e. RD1 and RD2), so that the expert can use any available contextual information.

Assume that a new road-piece instance, RP3, is created(Figure 13(b)). Then, the system provides this result to the instances involved in the situation, namely RD1 and RD2.

Suppose RD1 is the first to be informed of the proposed resolution. RD1 examines whether or not RP3 satisfies all constraints required to establish relation R1. In this case, however, RP3 fails, because RP3 is not adjacent to RP1. This failure activates an *exception handler*, which issues a top-down request to find a road-piece between RP1 and RP3(see Figure 13(c)).

Assume that another new road-piece instance, RP4, is detected(Figure 13(d)). Since RP4 is adjacent to RP1, RD1 establishes a PW relation to RP4, and then to RP3.

Figure 13(e) shows the data organization after the same analysis is performed by RD2. In this case, however, when RD2 establishes a PW relation to RP3, an exception handler in RP3 is triggered, because RP3 has two different

parents. More specifically, after RD2 establishes a PW relation to RP3, RD2 asks RP3 to check its reverse relation from RP3. An exception handler is activated as a result of this checking process. This handler issues a request to the system to examine the consistency between two parents. If they are consistent, the system merges the two PW hierarchies below them into one(Figure 13(f)). An exception handler of this kind is associated with each PW relation in order to construct a complete PW hierarchy by merging a pair of partial hierarchies.

There are several stages in the above example where the top-down request might have failed. In general, the Model Selection Expert has the ability to deal with such failures. Figure 14 shows a partial knowledge structure for suburban scenes. The Model Selection Expert analyzes the request to find RP3 (Figure 13(a)) by first assuming the road piece to be detected is a *visible road*, and issues a request to the Low Level Vision Expert. If this request fails, the Model Selection Expert switches to the other appearance of a road piece, i.e. an *occluded road*. The selection between *overpass* and *shadowed road* is done based on the cause of the failure. For example, if the cause of the failure is that the gray level in the overlapping region is too dark compared to the expected gray level, then the expert will hypothesize a *shadowed road*. If all efforts by the Model Selection Expert fail, this is reported to the High Level Expert. Then the system reports this to RD1 and RD2, which trigger their relevant exception handlers. Since different new hypotheses may be generated

by such exception handlers, no immediate further analysis is activated. Instead, these hypotheses are combined in the next interpretation cycle. In the case of Figure 13, RD1 and RD2 would both generate hypotheses for a road terminator.

If a top-down request issued by an instance fails, the instance activates another exception handler, if any. If all trials fail, the instance reports this to the system. Then the system activates another instance involved in the focused situation. The initial failure is not taken into account in any way by the system; this is a shortcoming of the present system.

## **1.2. Merging a Pair of Partial PW Hierarchies**

If a part instance is shared by two parent instances, the part issues a request to check the "similarity" between the parents. If they are similar, the system merges them into one.

Similarity examination involves checking whether or not the two parent instances denote(perhaps different pieces of) the same object. For example, RD1 and RD2 in Figure 13(e) should be merged into one road, although they do not denote the same (portion of) road. Knowledge about the continuity of roads is crucial in this example.

The more reliable of the two instances to be merged checks whether or not the *part instances* of the other instance are consistent with that more reliable parent. The more reliable parent may decide to merge with the other



parent, that such a merge is not (and will never be) possible (which places them in conflict) or that sufficient information is not available to make a decision.

Figure 15 illustrates an example of the third case. The definition of a house group is a group of regularly arranged houses which face the same side of the same road. As shown in Figure 15, if two house group instances share a house instance, the similarity examination is performed. If both house group instances face the same side of the same road instance, then they are similar and are merged into one. On the other hand, if one of them has not established such a "facing" relation, then it is not possible to verify the similarity between them. Moreover, even if the two house group instances have established "facing" relations to *different* road instances, it is still possible for them to be similar, because those road instances may be merged later. The house group instances can be regarded as conflicting only if their facing road instances are in conflict.

If the result of the similarity examination is "inconclusive", the system records the causes of the failure and *suspends* the action of establishing a new PW relation from a parent instance to the shared part instance. In the case shown in Figure 15, the relation between HG1 and H3 is suspended. The system records all suspended actions together with their causes. The suspended action can be reactivated if its cause is resolved by analyzing other situations.

## 6. Experimental Results

The image used in our experiment is a 320 by 160 portion of an aerial photograph(Figure 16) with intensities in the range of 0 to 63. The scene contains houses, roads, road intersections, trees, and driveways.

The appearance models are a subset of the possible models for suburban housing developments. Currently, we deal only with houses, road pieces, road intersections, and the spatial relations among them. Figure 14 shows the suburban housing development model used. In this section, we describe how our system proceeds to construct a road network interpretation from the image.

The system's analysis starts with a segmentation of the image. Since the houses and road pieces are modeled by compact and elongated rectangles, such rectangles are first extracted from the image. A simple blob finder and ribbon finder are used to find blobs and ribbons in the image.

Elongated rectangles are extracted and instantiated as road piece instances. These instances constitute the initial entries in the iconic database. Figures 16 shows the initial road-piece instances extracted from the image. As can be seen, roads are broken into pieces.

In the first cycle of the interpretation cycles, the system checks each instance and, for each relation, creates a hypothesis(for an SP relation or a top-down usage of a PW relation) or an instance(for a bottom-up usage of a

PW relation), if possible, and inserts it into the database. Since some of the relations may depend on yet undetermined values stored in frame slots, not all relations may be hypothesized at this point.

In the second cycle, the system's focus of attention mechanism selects the most promising situation. After a situation is selected, the system resolves it by first proposing a solution to it and then broadcasting messages to the source instances. Each source instance checks the proposed solution and requests the MSE to do top-down analysis if necessary. Also, the system may reorganize the database(e.g., unification of instances) during the resolution process.

In the current experiment, the MSE is simulated by a human. The descriptions of the action and the situation are displayed on the screen. The description of the result is entered from the terminal and is instantiated as an object instance and returned to the system.

Figures 17 - 23 show how the system proceeds to select a situation, resolve the selected situation, and reorganize the database as the result of resolving that situation. Figures 17 and 18 show two road-piece instances RP1, RP2, their parent instances RD1, RD2, and the hypotheses that RD1 and RD2 generate. During the hypothesis creation cycle, instances RD1 and RD2 create hypotheses H1, ... , H8. Hypotheses H4 and RP2 overlap(Figure 19.a). The system picks this situation(H4 and RP2 are consistent) and proceeds to resolve it.

Let C be the summarized constraints derived from the constraints of H4 and RP2. Since RP2 satisfies the constraint C, the system uses it as a proposed answer. RD1 checks the proposed solution, RP2, for adjacency. However, RP2 is not adjacent to RD1. RD1 issues a top down request to the MSE to find a road piece instance to connect RD1 and RP2. Currently, such a request is displayed on the screen and the result is entered from the terminal. The result can either be *success*, in which the description of the instance(object type and region description) is entered, or *failure*.

The description of a road piece instance(RP3) is entered from the terminal. MSE instantiates the instance and inserts it into the database. MSE reports RP3 to RD1. RD1 checks if RP3 is adjacent to RD1. Since RP3 is adjacent to RD1, RD1 establishes a PW link to RP3(Figure 20.b). Finally, RD1 checks MSG-A again and succeeds(since RD1 contains RP1 and RP3.) A PW link is established between RP2 and RD1(Figure 20.c). As a result, RP2 belongs to two parents. The system tries to unify them by checking if RD1 and RD2 are similar. In this case, they are similar. The system unifies RD1 and RD2 into a single instance(say RD'.) After the unification, road instance RD' has three parts(RP1, RP2, and RP3). Figure 21 shows the road instance RD' and its three parts. Figure 22 shows all the road instances after the selected situation is resolved.

During the unification process, several instances are merged into a single instance. The hypotheses generated by the merged instances are removed

from the database. A new set of hypotheses is generated in the next hypothesis creation cycle. Figure 23 shows the new hypotheses generated by RD'. Note that the original hypotheses H1, ... , H8 generated by RD1 and RD2 have been removed from the database.

Figure 24 shows a case where alternate hypotheses are generated. A road can either be extended continuously, or stop at a road terminator. One way to conduct the search is to look for the adjacent road piece first. If that search fails, then the search for a road terminator can start. Such a strategy is illustrated in Figure 24.a. Figure 24.b shows a road instance and the alternate hypotheses it generates during the process.

Figure 25.a shows the final result of constructing the road network interpretations by the system. The interpretation graphs are shown in Figure 25.b. Each node represents an instance. There are 29 road piece instances, 10 road instances, and 5 road terminator instances. Figure 26 shows the road joint instance J1 and all road instances meeting there. Figure 27 shows road instance R2, the road terminator instances adjacent to it, and its part objects.

## REFERENCES

- [1] V. Hwang, T. Matsuyama, L.S. Davis, and A. Rosenfeld, "Evidence Accumulation for Spatial Reasoning in Aerial Image Understanding", Technical Report CAR-TR-28, Center for Automation Research, University of Maryland, 1983.
- [2] P.G. Selfridge, "Reasoning about Success and Failure in Aerial Image Understanding", Ph. D. thesis, University of Rochester, 1982.
- [3] M. Minsky, "A Framework for Representing Knowledge", in *The Psychology of Computer Vision* (P. Winston, ed.), McGraw-Hill, NY, 1975.
- [4] R.L. Haar, "The Representation and Manipulation of Position Information Using Spatial Relations", Technical Report TR-923, Computer Vision Laboratory, University of Maryland, 1980.
- [5] D. McDermott, "A Theory of Metric Spatial Inference", *Proc. of Natl. Artificial Intelligence Conference*, pp.246-248, Aug. 1980.

Check the validity of  $REL(O1, O2)$  for a pair of pictorial entities which may be instances of  $O1$  and  $O2$ .

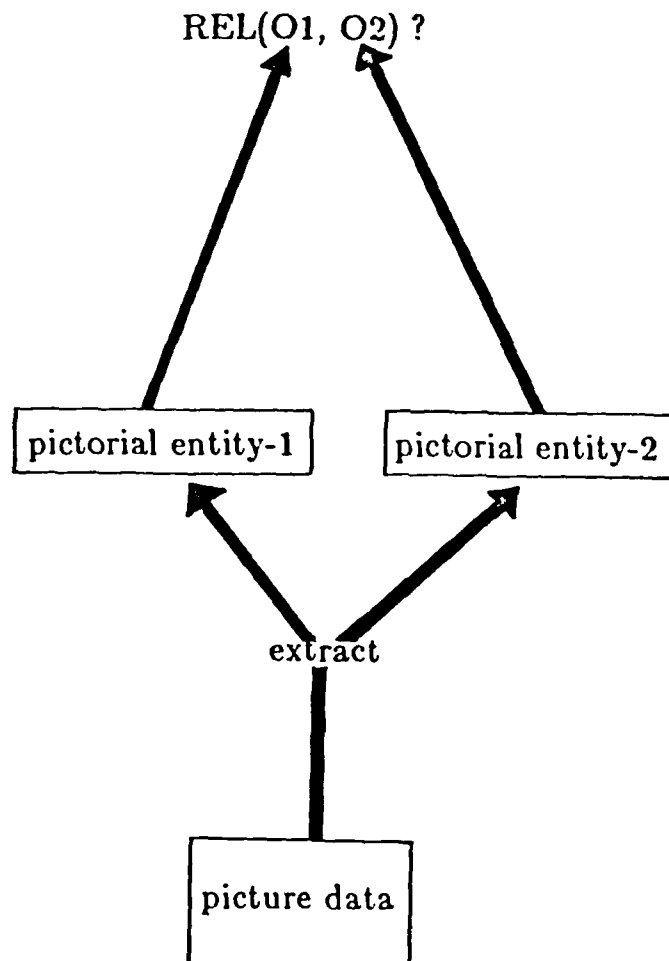


Fig. 1 Using a relation as a constraint.

## Hypothesis Generation

instance of object 02

hypothesis for object 01

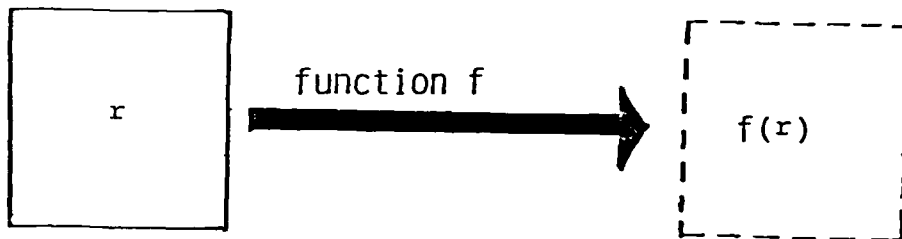


Fig. 2 Hypothesis generation based on functional representation of a relation



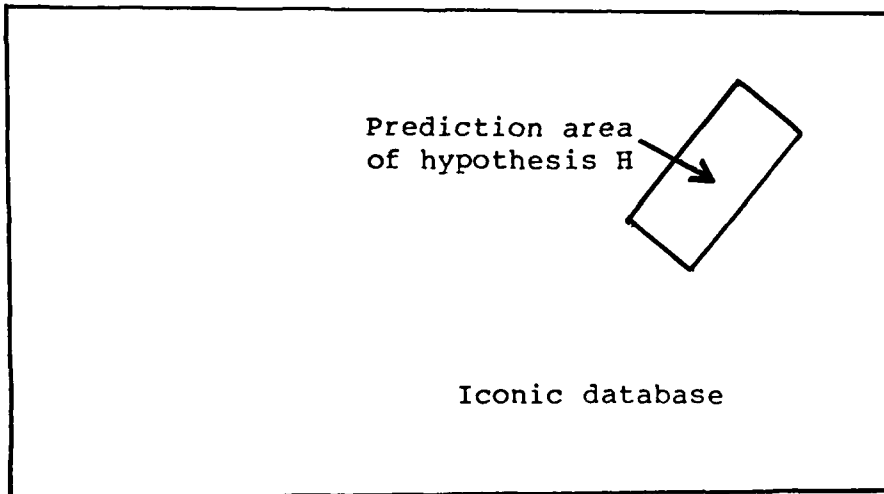
Frame name : Road piece  
Slot name : Length  
            Width  
            Direction  
            Coordinate of the local coordinate system  
            Father

(1) The description of the road piece frame

Frame name : Road  
Slot name : Total-length  
            Average-direction  
            Left-adjacent-road-piece  
            Right-adjacent-road-piece  
            Left-connecting-road-terminator  
            Right-connecting-road-terminator  
            Left-neighboring-house-group  
            Right-neighboring-house-group

(2) The description of the road frame

Figure 3 : (a) The description of the road frame and the  
            road piece frame.



(1) Iconic description of hypothesis H

```
(AND (EQUAL OBJECT-TYPE ROAD)
      (AND (LESSP TOTAL-LENGTH 100)
            (GREATERP TOTAL-LENGTH 50))
      (AND (LESSP AVERAGE-WIDTH 15)
            (GREATERP AVERAGE-WIDTH 10))
      (AND (LESSP AVERAGE-DIRECTION 50)
            (GREATERP AVERAGE-DIRECTION 30)))
```

(2) Symbolic description of hypothesis H

Figure 3 : (b) The description of a road hypothesis H.

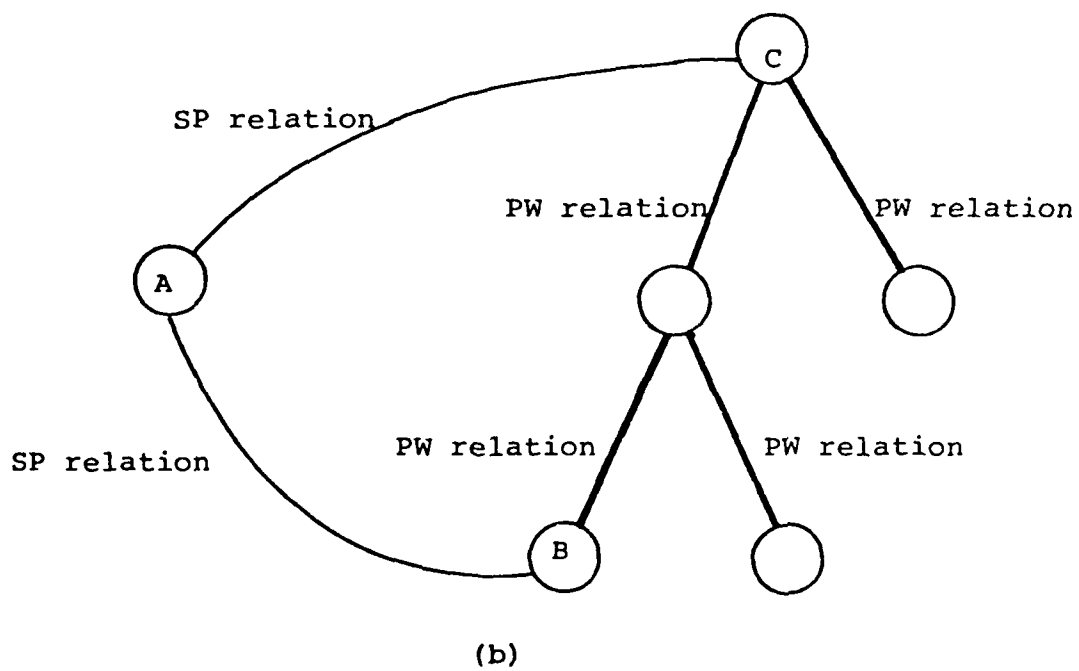
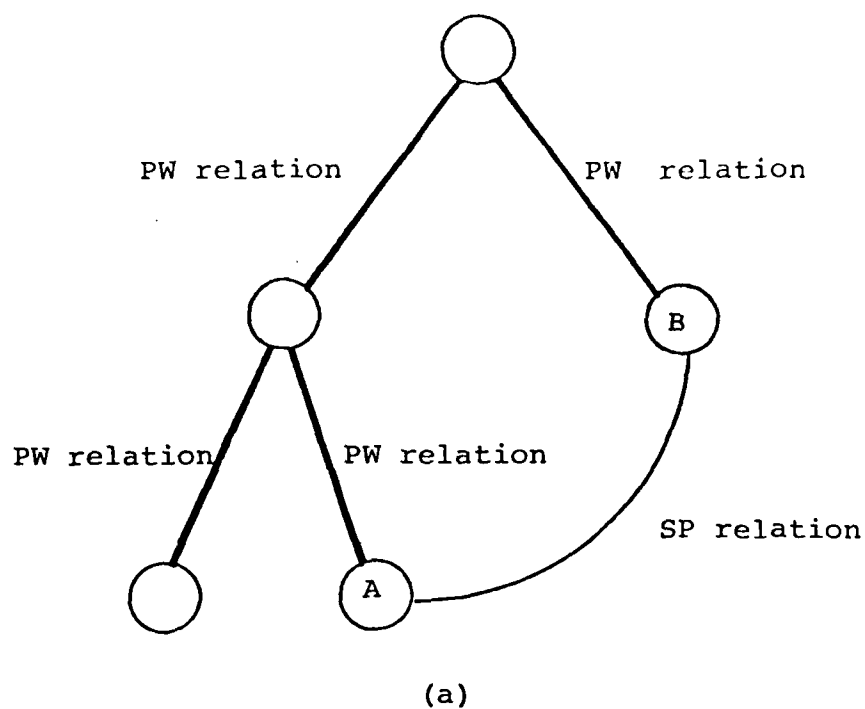
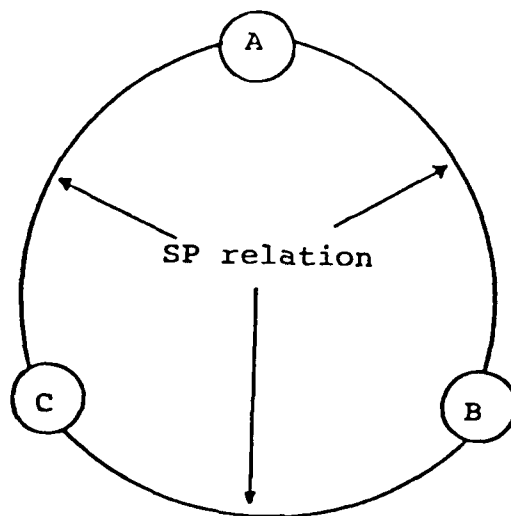


Fig. 4 Avoiding redundant accumulation of evidence



(c)

Figure 4 (cont.)

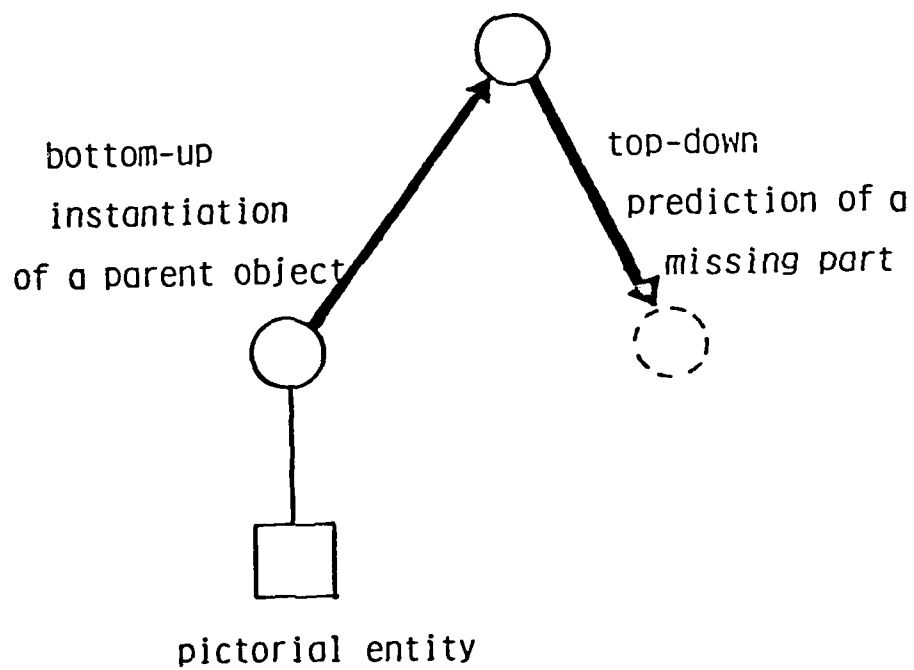


Fig. 5 Hypothesis generation by a part-whole relation

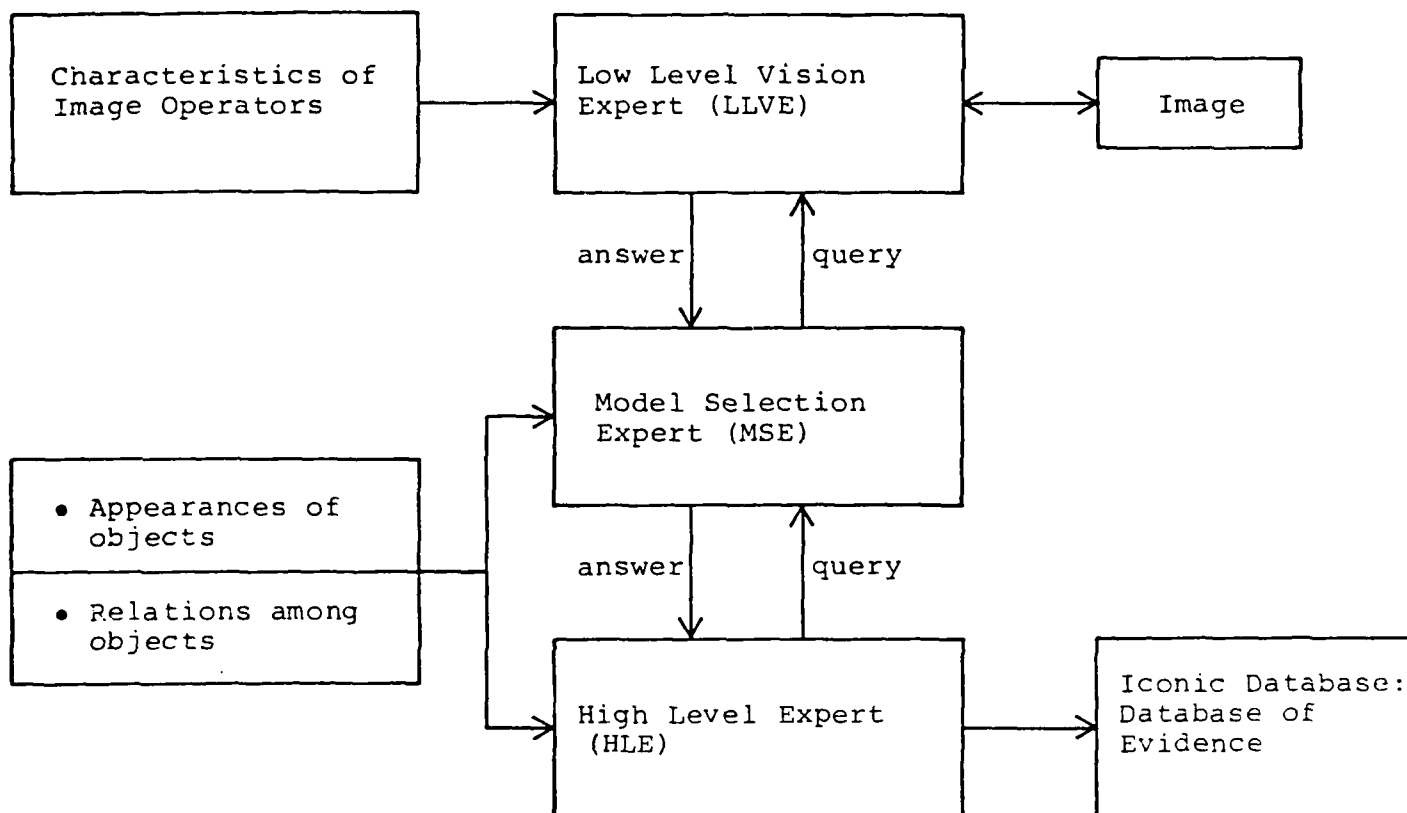
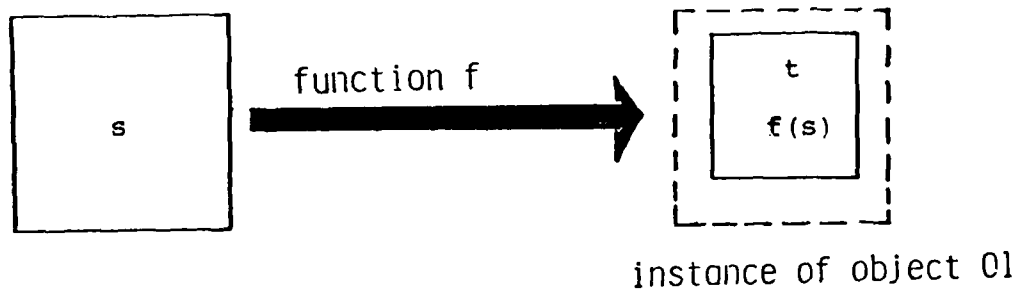


Figure 6 An Image Understanding System

instance of object 02

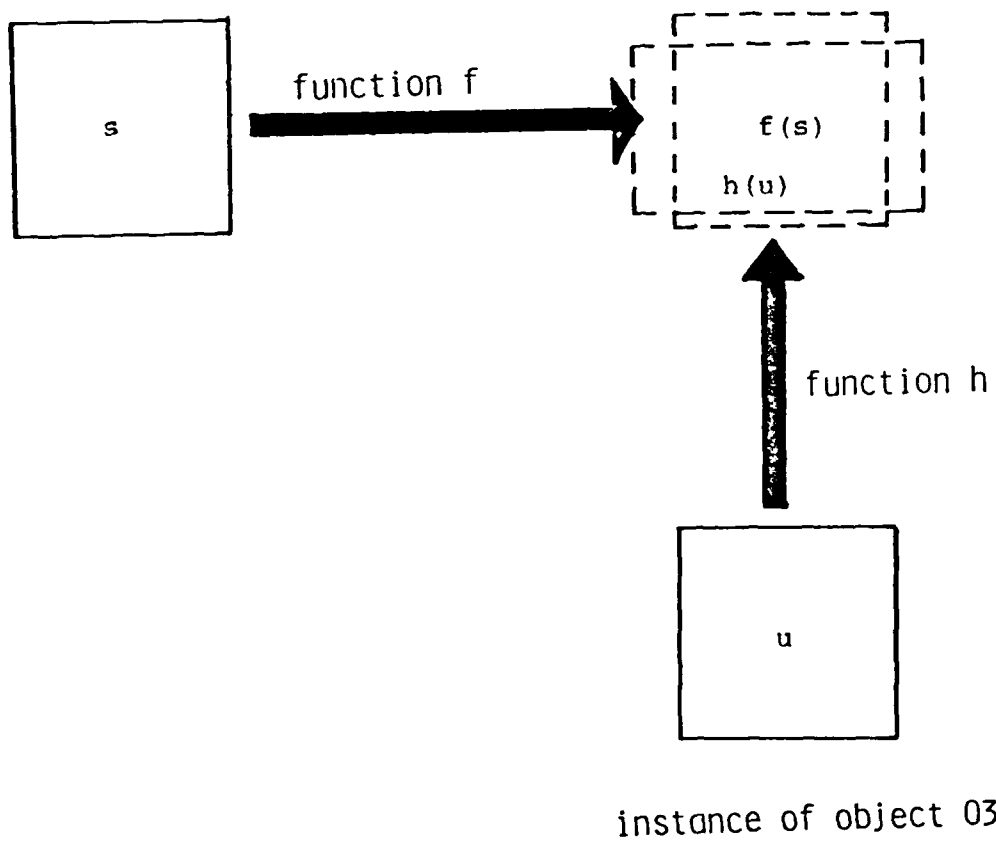
hypothesis for object 01



(a)

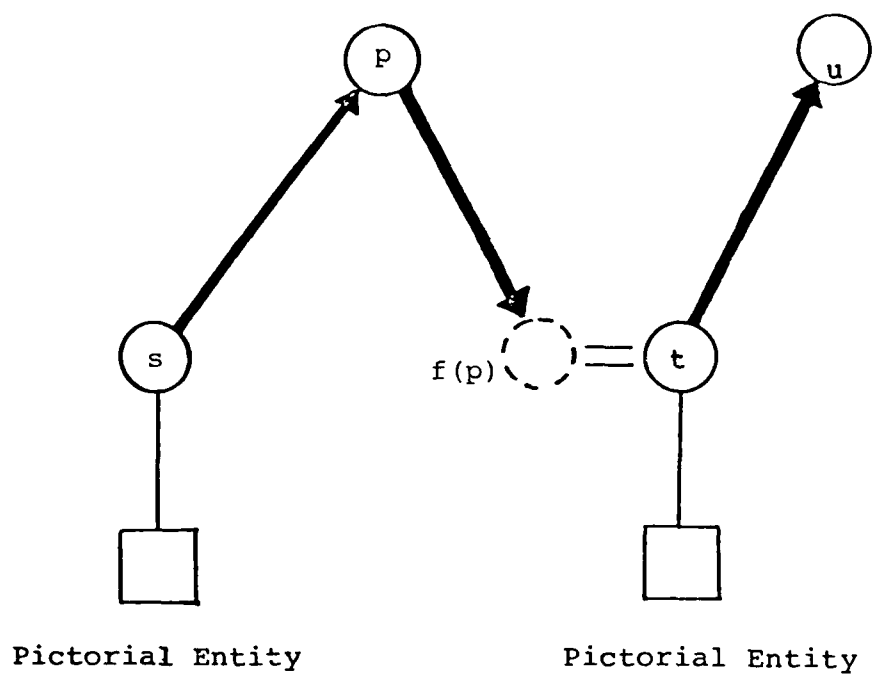
instance of object 02

hypothesis for object 01

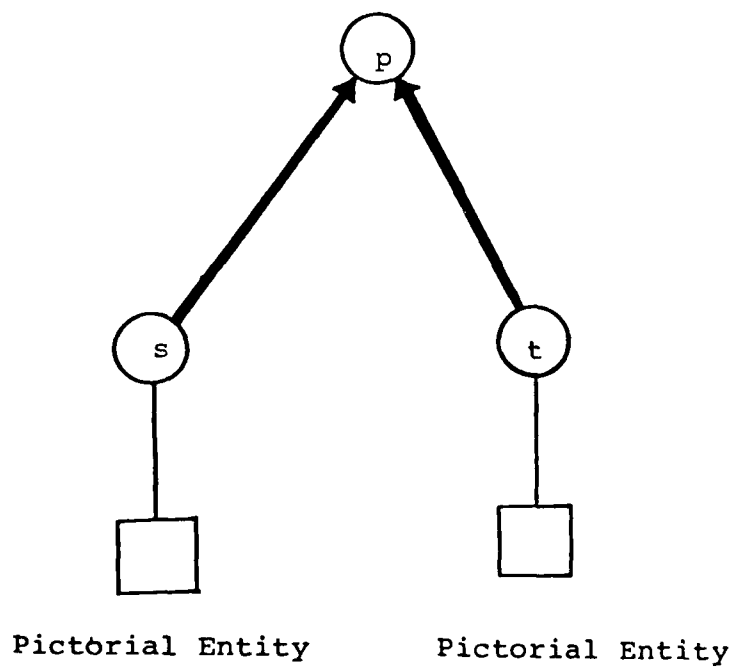


(b)

Fig. 7 Principle of evidence accumulation for spatial reasoning



(a)



(b)

Fig. 8 Constructing a part-whole hierarchy



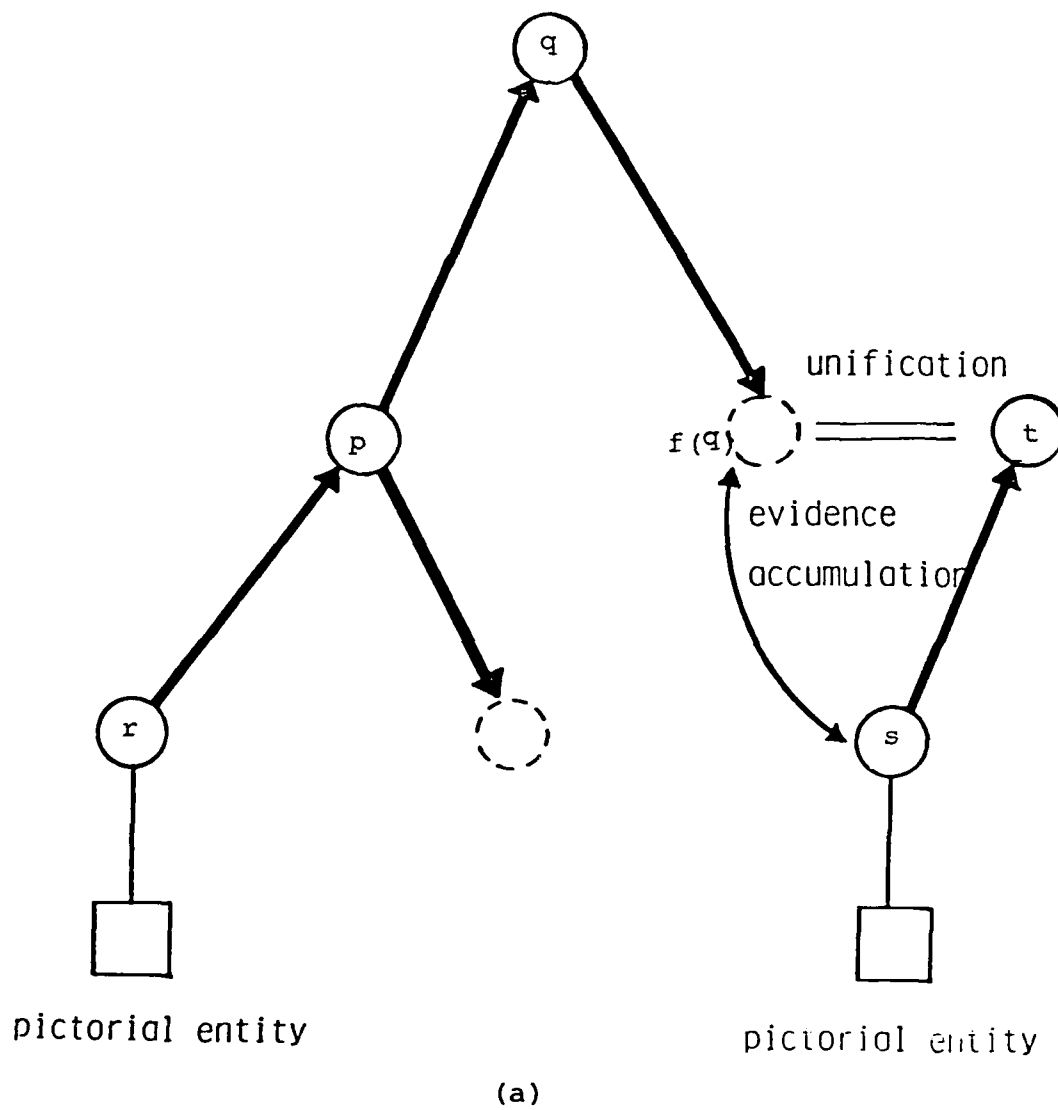


Fig. 9 (a) Another example of constructing a part-whole hierarchy

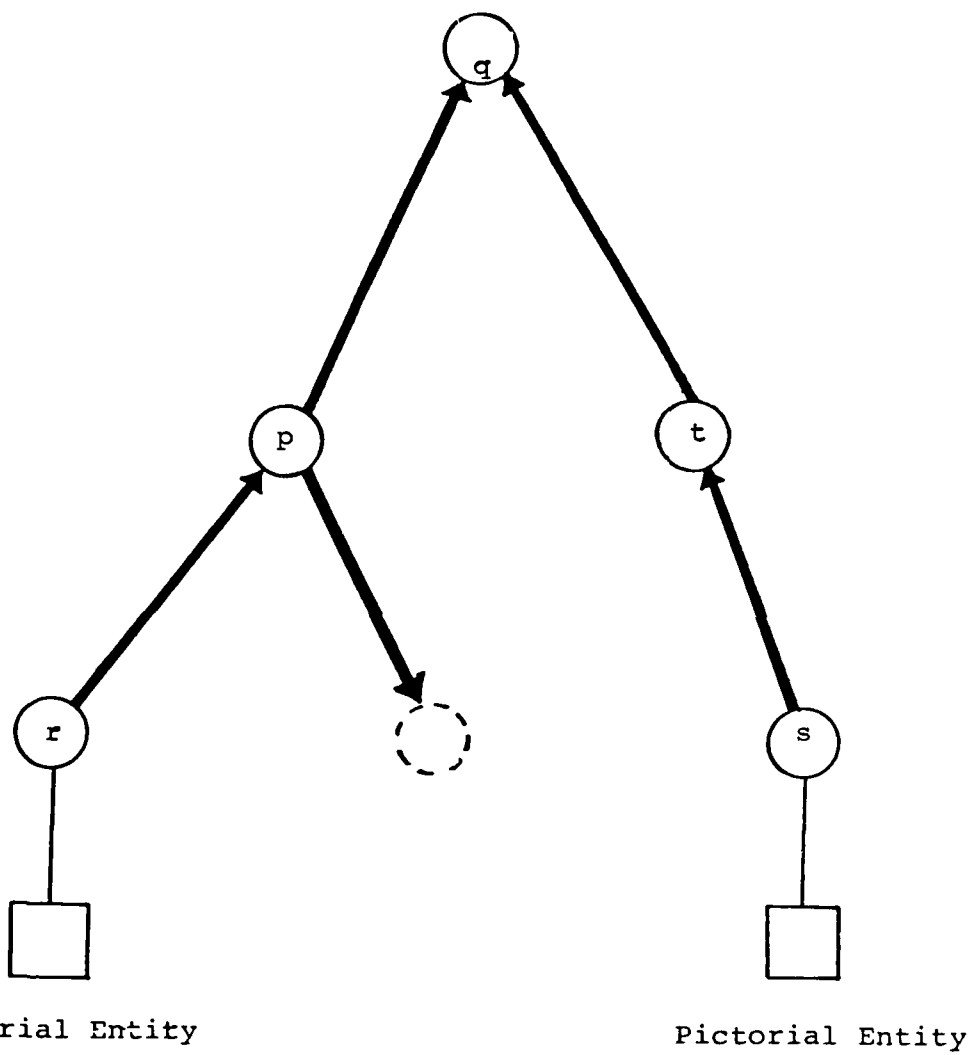


Fig. 9 (b) Result of the unification

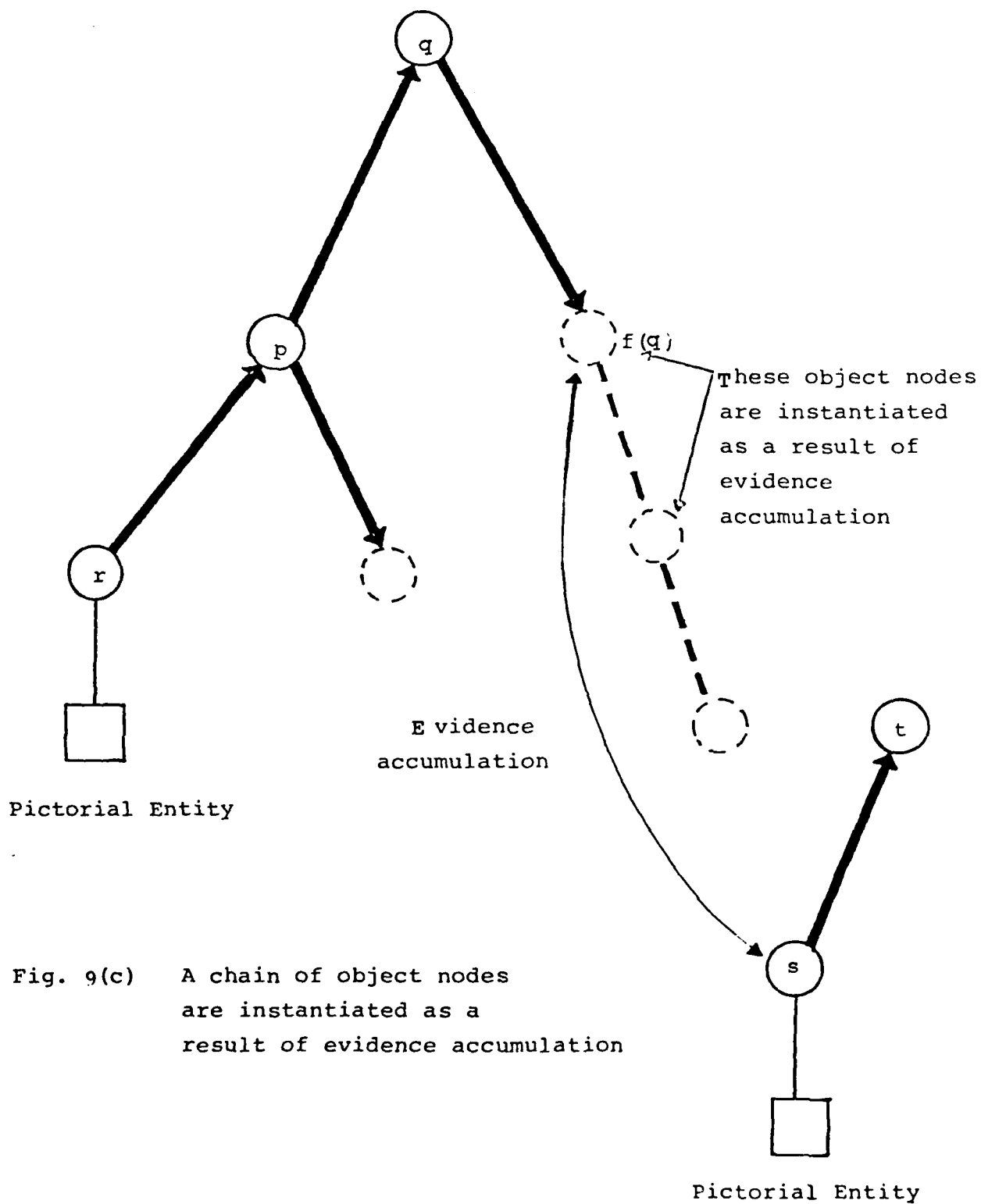
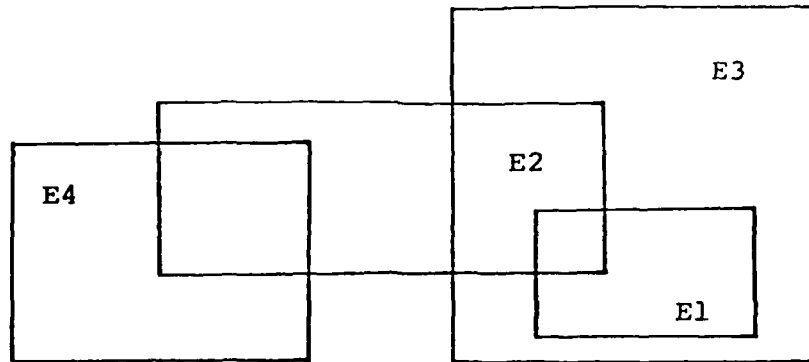
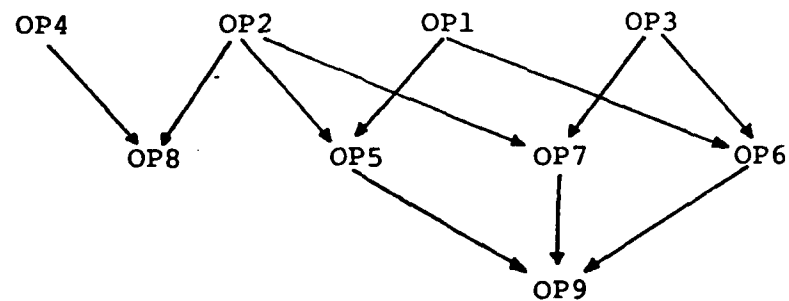


Fig. 9(c) A chain of object nodes are instantiated as a result of evidence accumulation



(a) Four overlapping pieces of evidence

Overlap	Constituent Evidence
OP1	E1
OP2	E2
OP3	E3
OP4	E4
OP5	E1, E2
OP6	E1, E3
OP7	E2, E3
OP8	E2, E4
OP9	E1, E2, E3



(b)

Fig. 10 Lattice structure to represent overlaps among pieces of evidence

```

(AND (EQUAL OBJECT-TYPE ROAD-PIECE)
      (AND (LESSP LENGTH 19)
            (GREATERP LENGTH 14))
      (AND (LESSP DIRECTION 60)
            (GREATERP DIRECTION 45)))

```

(a) The description of constraint C1.

Slot name translation table

Slot name of road-piece frame	Slot name of road frame
Length	Total-length
Width	Average-width
Direction	Average-direction

(b) Slot name translation table for the PW relation between the road frame and the road piece frame.

```

(AND (EQUAL OBJECT-TYPE ROAD)
      (AND (LESSP AVERAGE-LENGTH 19)
            (GREATERP AVERAGE-LENGTH 14))
      (AND (LESSP AVERAGE-DIRECTION 60)
            (GREATERP AVERAGE-DIRECTION 45)))

```

(c) The description of constraints C1 after translation.

Figure 11 : Translation of constraints.

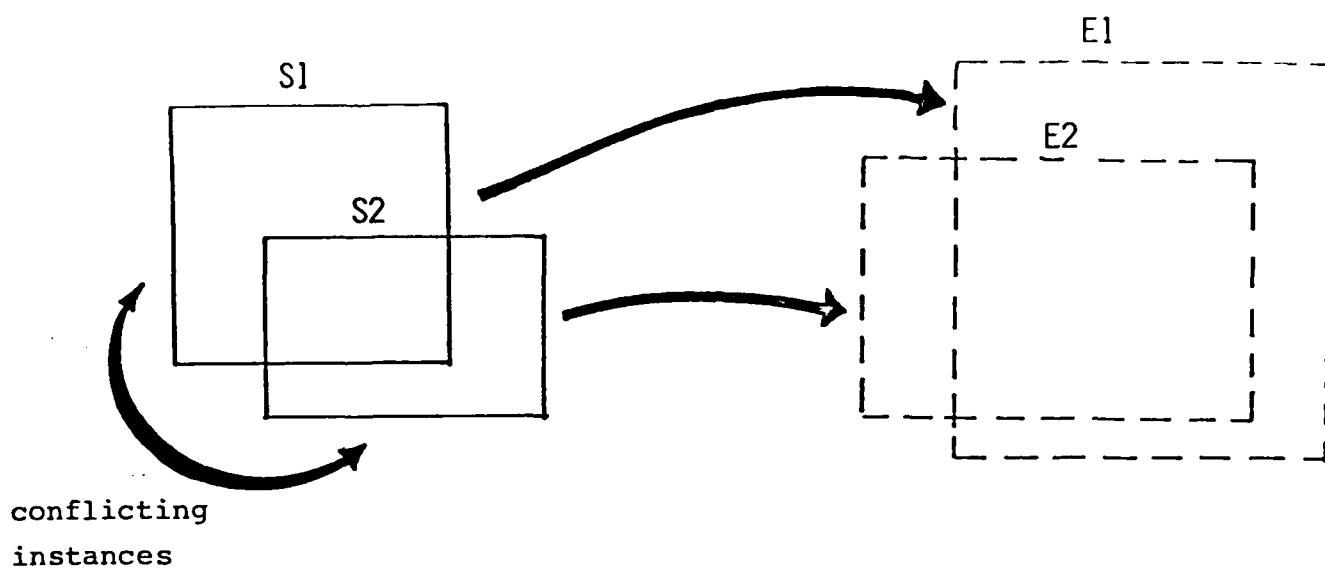
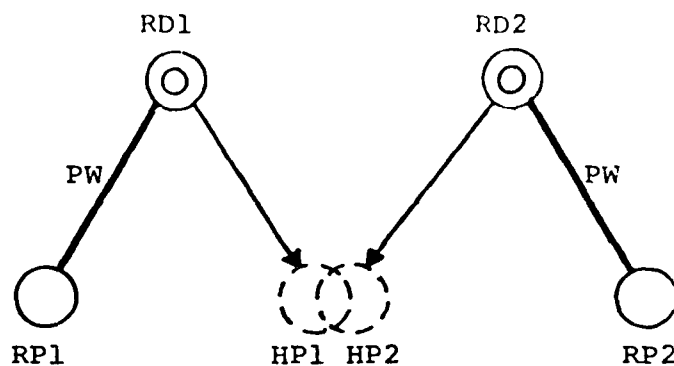
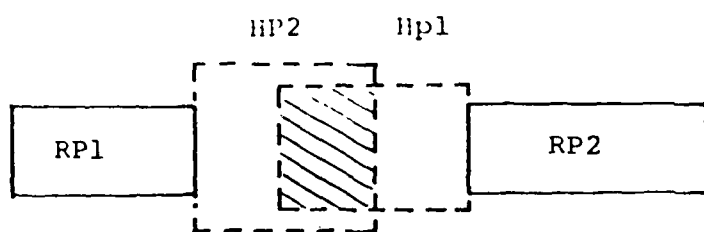
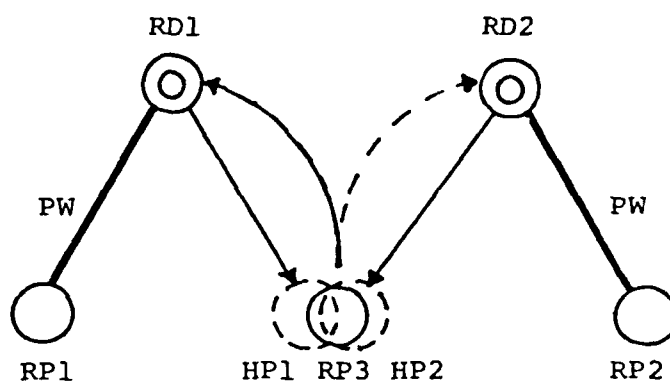
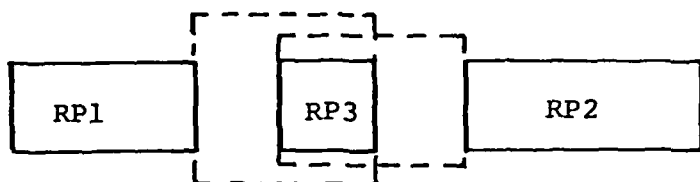


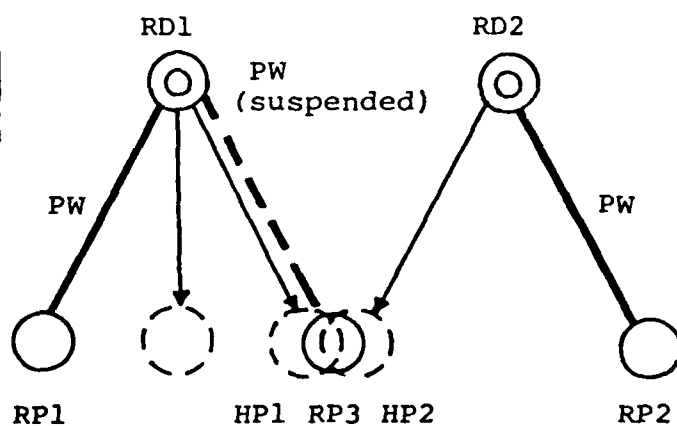
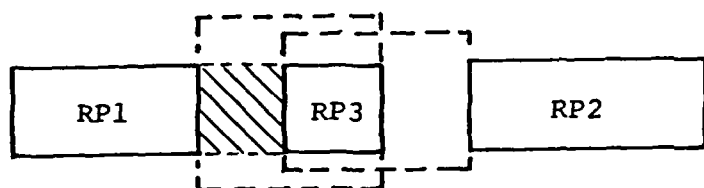
Fig. 12 Hypotheses generated by conflicting instances



(a)

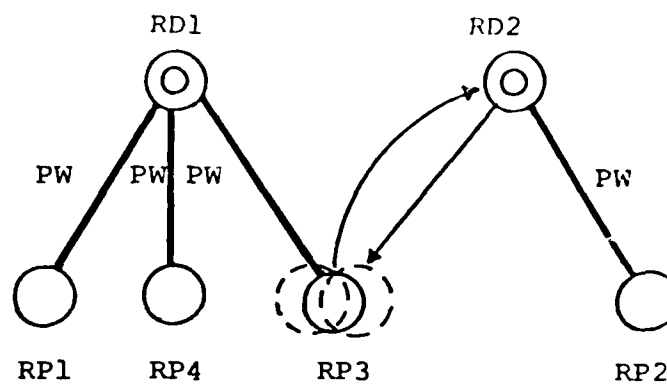
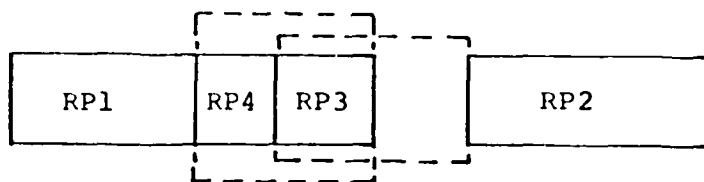


(b)

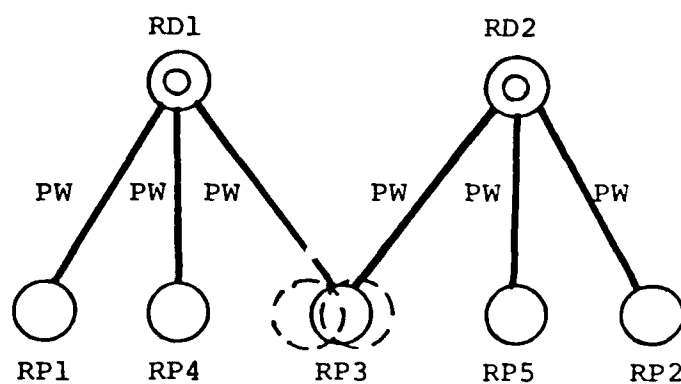
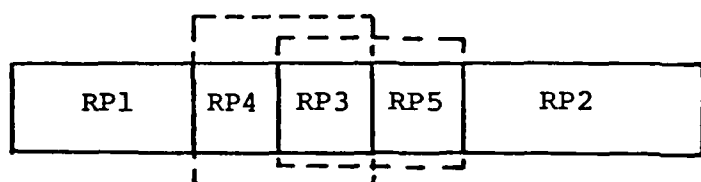


(c)

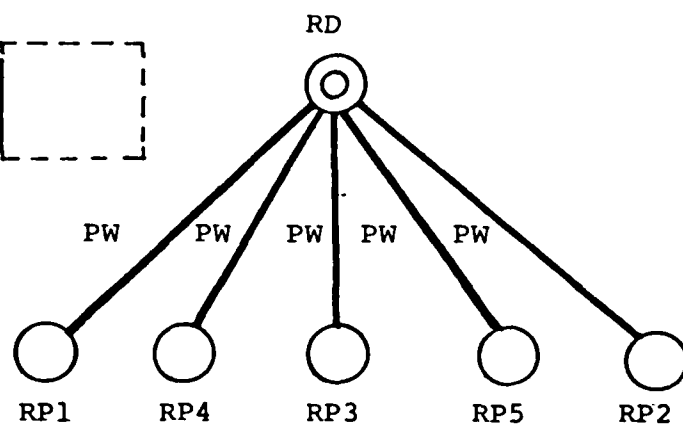
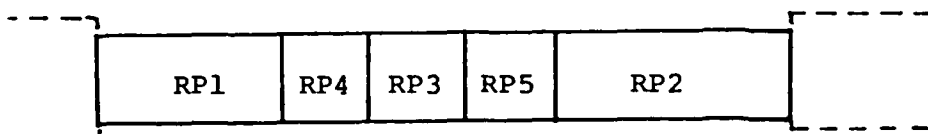
Figure 13 Resolving a situation(an example)



(d)



(e)



(f)

Fig. 13 Resolving a situation (see text)



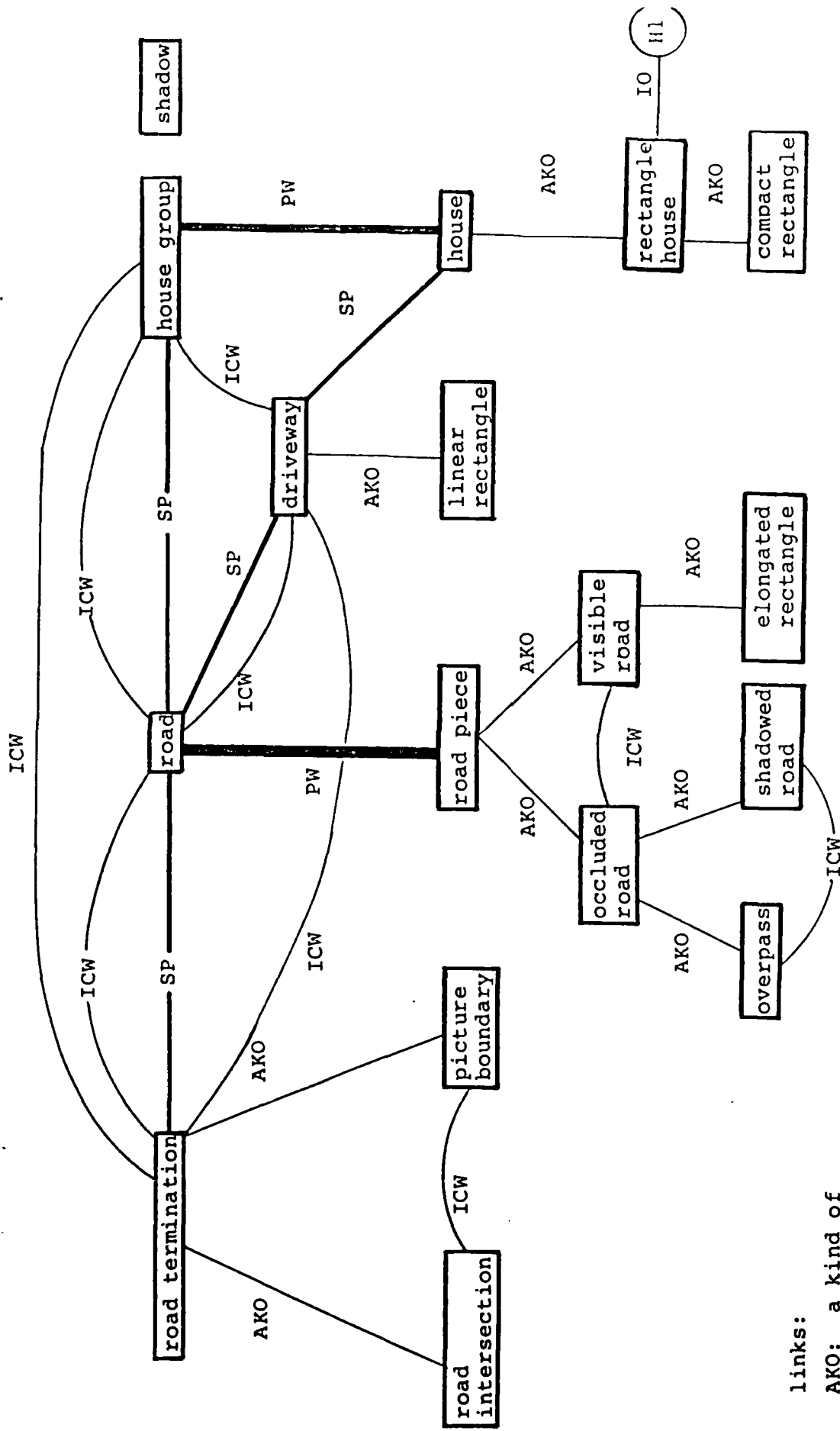


Fig. 14 Knowledge organization about suburban scenes

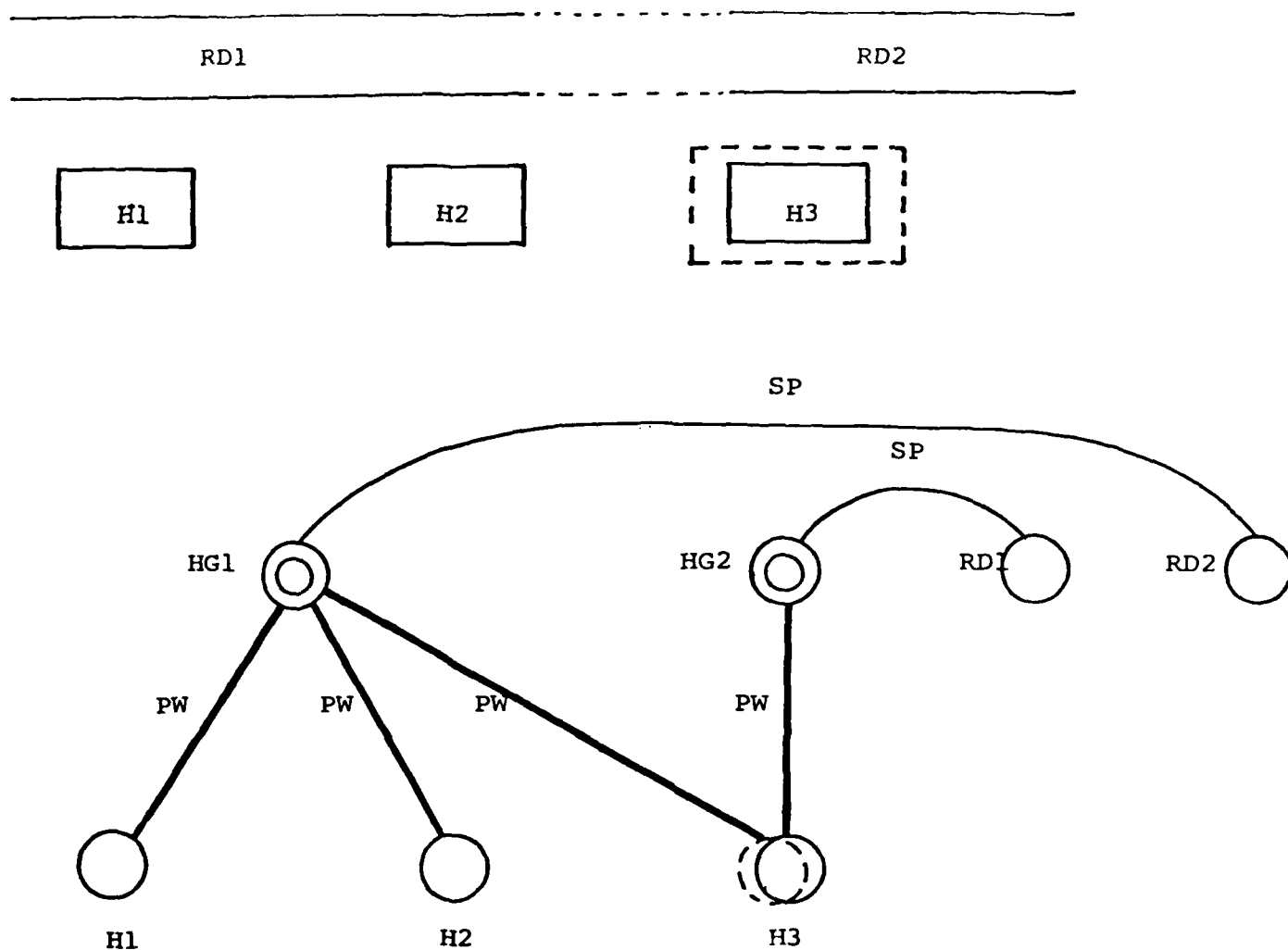


Fig. 15 Merging two partial PW hierarchies

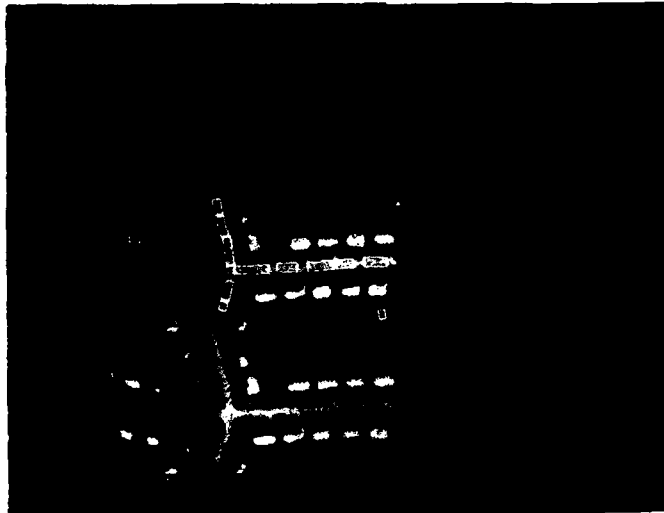
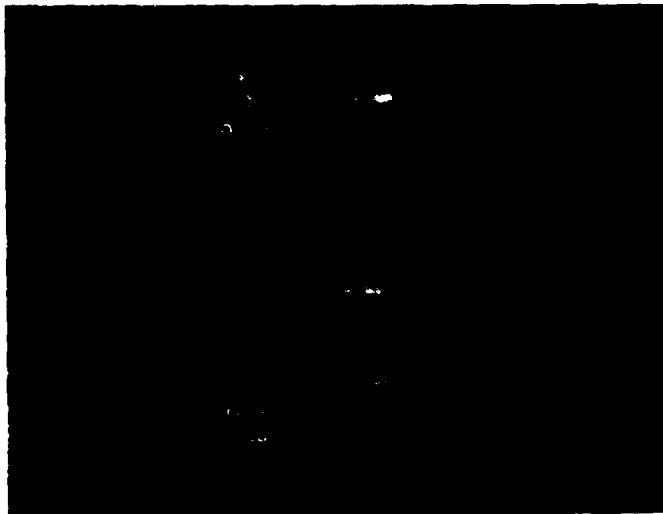
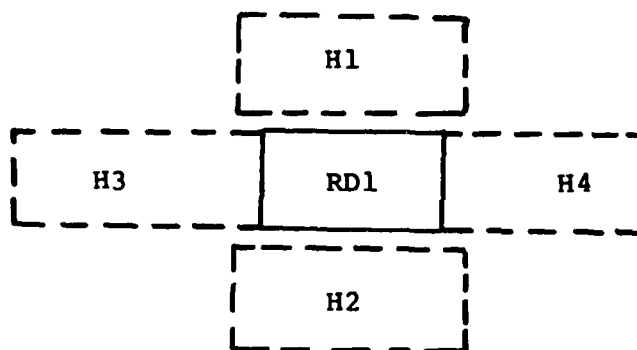


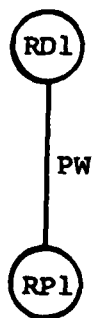
Figure 16 : Original image(bottom) and initial road piece instances overlapped on the image(top).



(a) A road instance RD1(bottom), the neighboring house group hypotheses(H1, H2) (middle), and the adjacent road piece hypotheses(H3, H4).

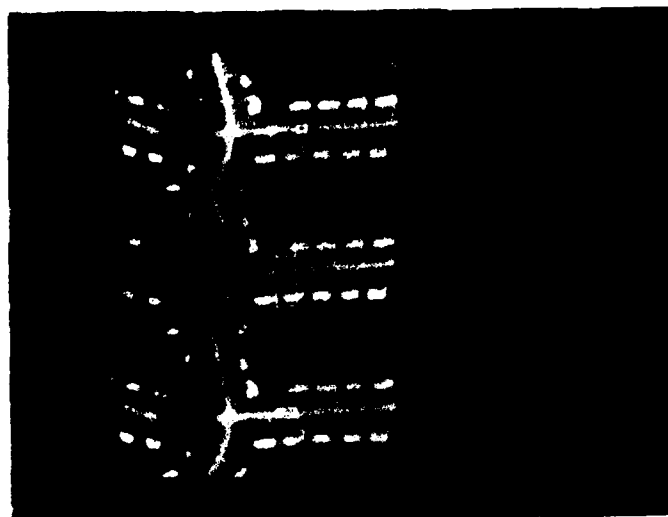


(b) A depiction of RD1 and the hypotheses it generates

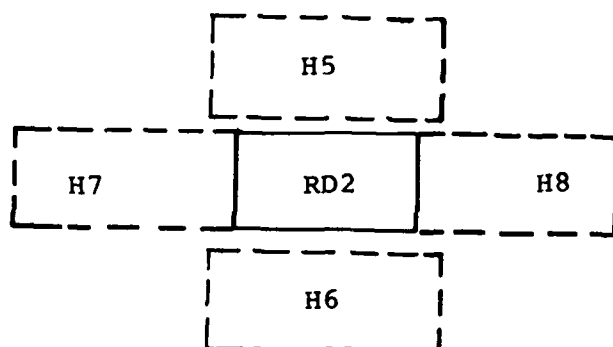


(c) The interpretation graph of RD1.

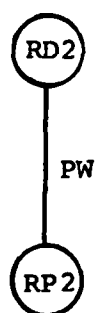
Figure 17 : A road piece instance RP1, its parent RD1, and the hypotheses RD1 generates.



(a) A road instance RD2(bottom), the neighboring house group hypotheses(H5, H6) (middle), and the adjacent road piece hypotheses(H7, H8).

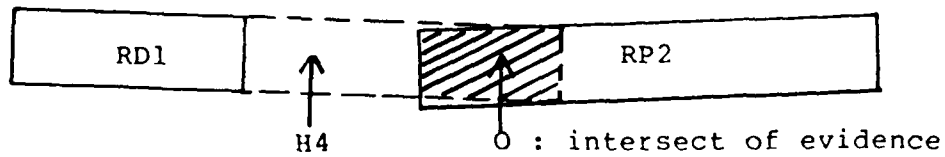


(b) A depiction of RD1 and the hypotheses it generates

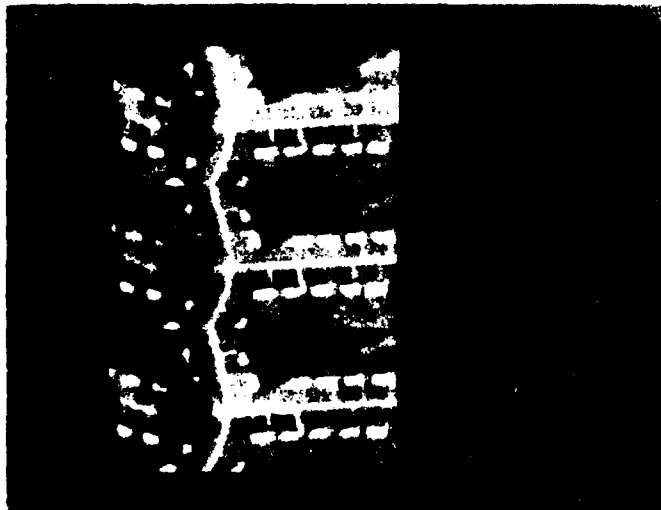


(c) The interpretation graph of RD2.

Figure 18 : A road piece instance RP2, its parent RD2, and the hypotheses RD2 generates.



(a) A depiction of the situation

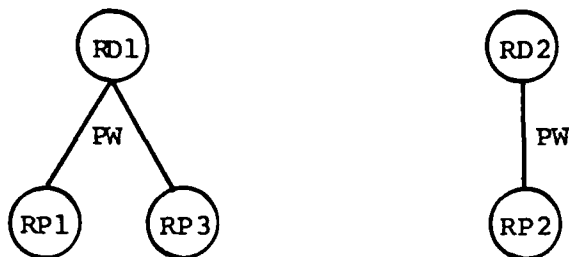


(b) The supporting sources of the situation(bottom), the region of top-down prediction request(middle), the road piece instance entered from the terminal(top).

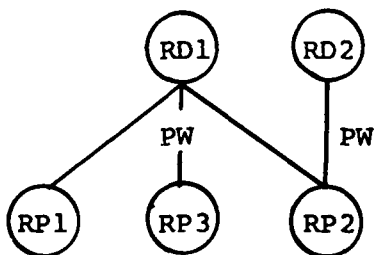
Figure 19 : A situation



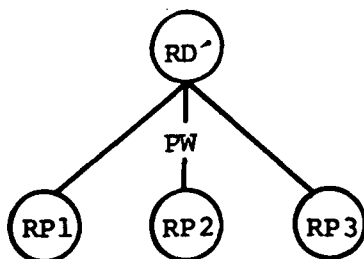
(a) The interpretation graphs before resolving the situation.



(b) The interpretation graphs after road piece RP3 is entered into the iconic database.



(c) The interpretation graph after RD1 rechecks its message.



(d) The interpretation graph after the unification of RD1 and RD2.

Figure 20 : The interpretation graphs during the resolution of a situation.

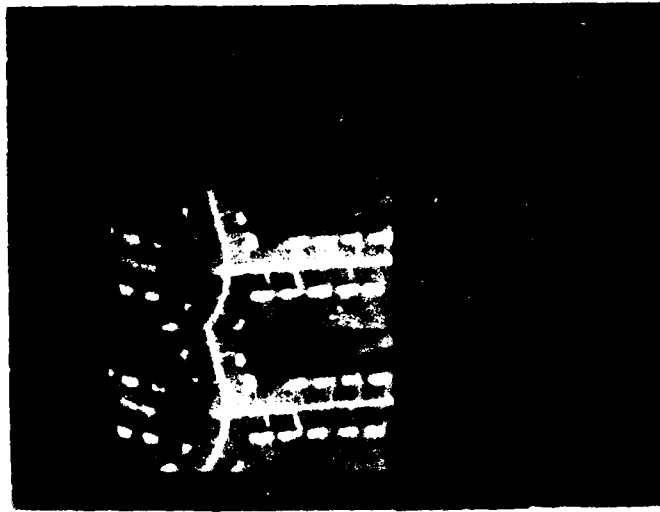


Figure 21 : Resolving a situation. Road instance RD3 (bottom) and its part objects (RP1, RP2, and RP3) (top).



Figure 22 : All road instances after the situation is resolved.



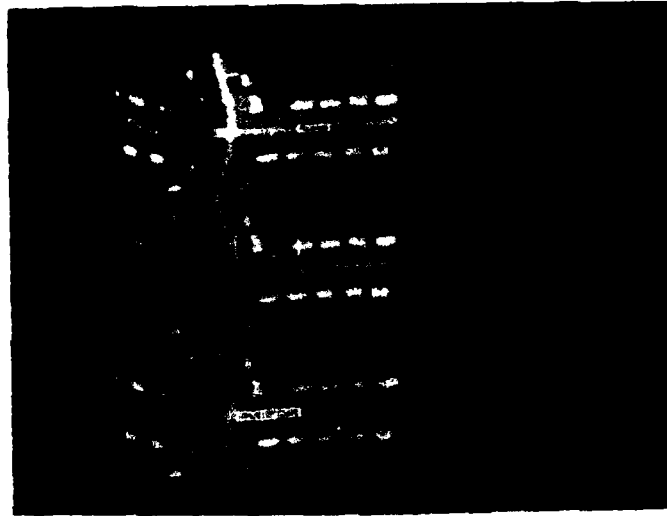
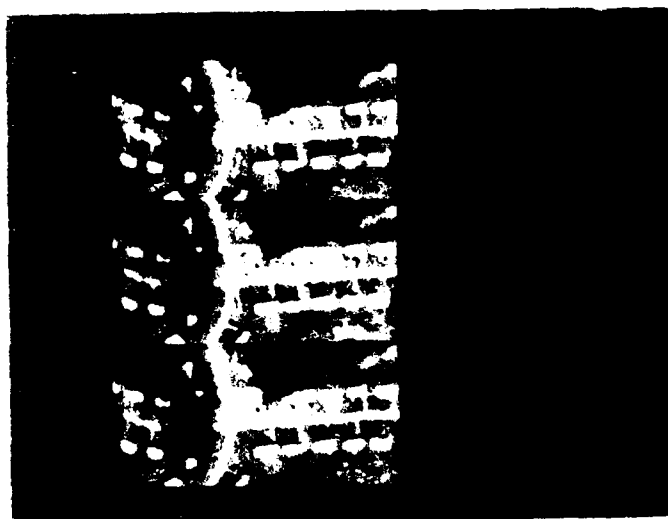


Figure 23 : Update of hypotheses : road instance RD3(bottom)  
neighboring house group hypotheses(middle), and  
adjacent road piece hypotheses(top).

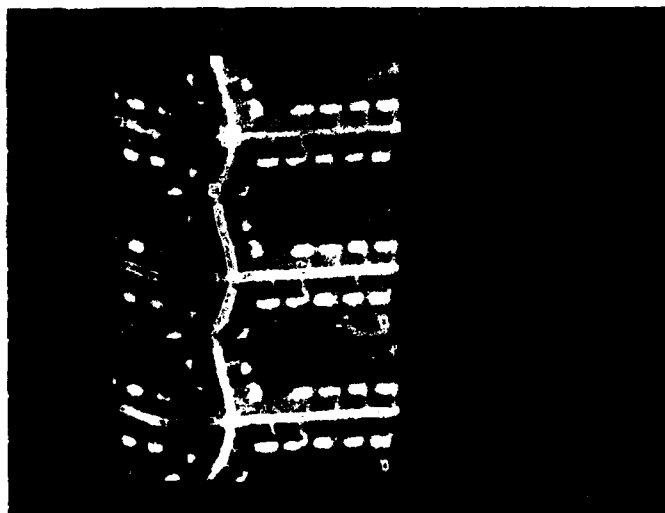
Relation	Precondition
Adjacent Road piece	Always
Adjacent Road terminator	When search for adjacent road piece has failed

- (a) Precondition for adjacent road piece and adjacent road terminator relations.



- (b) A road instance(bottom), adjacent road piece hypotheses(middle), and adjacent road terminator hypothesis(top).

Figure 24 : Change of hypotheses.



(a) All road piece instances(bottom), all road instances(middle), and all road terminator instances(top).

Figure 25 : Final interpretation of the road network.

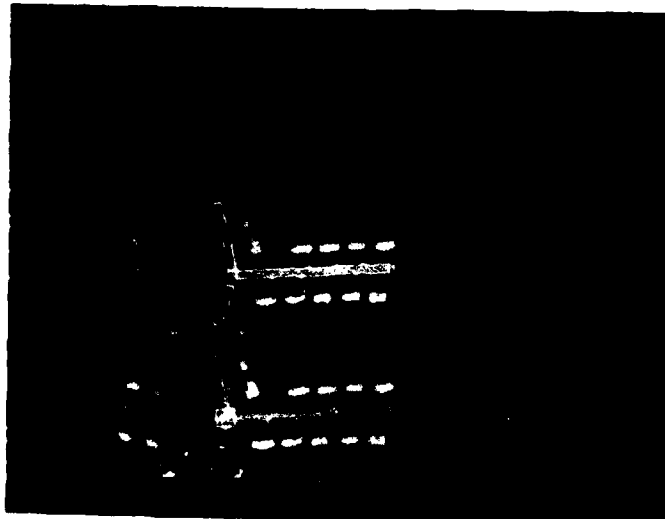


Figure 26 : Road joint instance J1(bottom) and all road instances intersecting at J1(top).

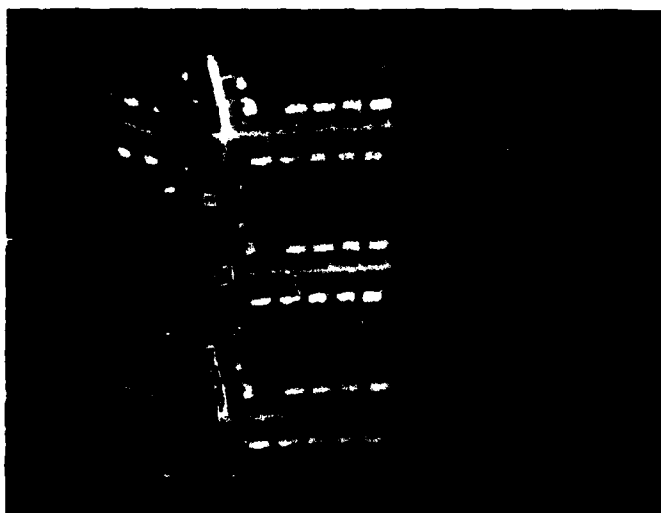
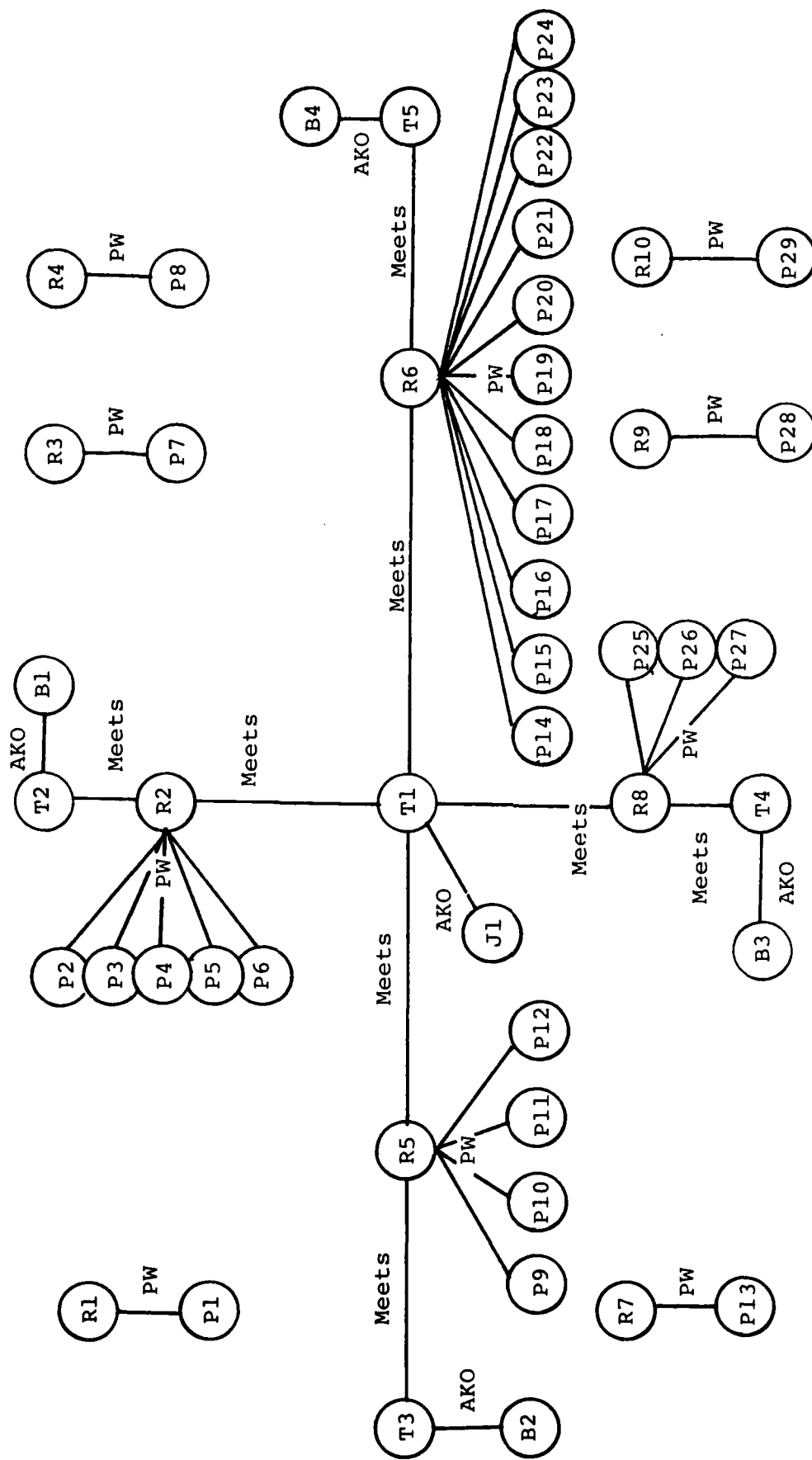


Figure 27 : Road instance R2(bottom), road terminator instance instances adjacent to it(middle), and road piece instances contained in it(top).



Legend : R - Road instance  
 P - Road piece instance  
 T - Road terminator instance  
 B - Picture boundary instance  
 J - Road joint instance

Figure 25 (b) The interpretation graph constructed by the system.

## UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
	AD 4142 174	
4. TITLE (and Subtitle) EVIDENCE ACCUMULATION FOR SPATIAL REASONING		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER CAR-TR-54; CS-TR-1381
7. AUTHOR(s) Takashi Matsuyama Vincent Shang-Shouq Hwang Larry S. Davis		8. CONTRACT OR GRANT NUMBER(s) F49620-83-C-0082
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Automation Research University of Maryland College Park, MD 20742		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A2
11. CONTROLLING OFFICE NAME AND ADDRESS Math. & Info. Sciences, SGODT/NM Bolling AFB Washington, DC 20332		12. REPORT DATE March 1984
		13. NUMBER OF PAGES 58
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Image understanding Aerial photographs Spatial relations		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  This paper describes the evidence accumulation process of an image understanding system first described in [1], which enables the system to perform top-down (goal-oriented) picture processing as well as bottom-up verification of consistent spatial relations among objects.		